Available online at www.sciencedirect.com
ScienceDirect
Control Engineering Practice 16 (2008) 930–945
CONTROL ENGINEERING PRACTICE
www.elsevier.com/locate/conengprac

Discrete linear control enhanced by adaptive neural networks in application to a HDD-servo-system ☆

Guido Herrmann[a,*,1], Shuzhi Sam Ge[b], Guoxiao Guo[c,2]

[a]Dynamics and Control Research Group, University of Bristol, University Walk, Bristol BS8 1TR, UK
[b]Department of Electrical and Computer Engineering, National University of Singapore, Singapore 119260, Singapore
[c]Western Digital, Lake Forest, CA 92630-7741, USA

Received 2 October 2005; accepted 29 October 2007
Available online 20 December 2007

Abstract

The performance of a linear, discrete high performance track following controller in a hard disk drive is improved for its disturbance rejection by the introduction of a discrete non-linear, adaptive neural network (NN) element. The NN-element is deemed to be particularly effective for rejection of bias forces and friction. Theoretical, simulation and experimental results have been obtained. It is shown theoretically that an NN-element is effective in counteracting these non-linear, system-specific, model-dependent disturbances. The disturbance, i.e. the bias and friction force, is assumed to be unknown, with the exception that the disturbance is known to be matched to the plant actuator input range and the disturbance is an (unknown) continuous function of the plant output measurements. For a non-linear simulation model and a laboratory HDD-servo-system, it is shown that the NN-control element improves performance and appears particularly effective for a reasonably small number of NN-nodes.

Keywords: Hard-disk; Servo-control; Bias and friction forces; Compensation; Neural networks

1. Introduction

One significant development has characterized the technological advancement of hard disk drives (HDDs): the increase in track density which has now reached about 150 kTPI (TPI—track per inch) in consumer products and 400 kTPI on a high performance spin stand in a laboratory environment (Du et al., 2006). This development has been driven by consumer demand for higher storage capacity of hard disks and a more varied application area for HDDs, which require a smaller form factor (e.g. application in portable equipment such as cameras) or large storage space (e.g. modern 'video' recorders storing digital movies). The demand for increased data density and a greater spectrum of HDD application areas implies the need of significant advancements of the overall HDD-technology which allows to convert novel ideas as for the high performance 400 kTPI-spin stand (Du et al., 2006) into marketable low-cost products. One way of approaching this is to introduce dual-stage actuator technology (e.g. Herrmann & Guo, 2004; Horowitz, Li, Oldham, Kon, & Huang, 2007) or by using highly accurate mechanics for the low-cost Voice Coil Motor (VCM)-actuator, which is important for small form-factor drives and a decreased head-disk spacing around 10 nm. Furthermore, the VCM-arm has to be constructed to minimize pivot bearing friction and other bias forces on the arm. Despite the advances in HDD-technology, these forces are known to pose a significant problem in future high density HDDs (Gong, Guo, Lee, & Bin, 2002).

In terms of servo-control, the VCM-actuator has to deal with the following servo-tasks: seek/settling and track following. Seek/settling control has to ensure a fast

☆This work has been partially supported by the A*Star SERC Grant no. 052 101 0097.
*Corresponding author. Tel.: +44 117 928 8100; fax: +44 117 929 4423.
E-mail address: g.herrmann@bris.ac.uk (G. Herrmann).
[1]This work was completed during a two year Senior Research Fellowship of Guido Herrmann at the A*Star, Data Storage Institute, Singapore.
[2]The author was with the A*Star Data Storage Institute, Singapore, when this work was carried out.

0967-0661/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.conengprac.2007.10.011

movement of the read/write head from one track to another using a non-linear time-optimal controller. For track following, linear high bandwidth controllers are necessary to ensure good error rejection capabilities for counteracting disturbances such as pivot bearing friction which are particularly detrimental to track following and short span track seeking. These servo-control tasks are only achieved using advanced sampled-data/discrete control methods, as a HDD-servo-system is inherently a sampled-data control system due to the way servo/position-information is stored on a hard disk (Abramovitch & Franklin, 2002; Al Mamun & Ge, 2005).

It has been acknowledged that the modelling task for bias forces and for friction is complex (Hensen, van de Molengraft, & Steinbuch, 2002; Oboe, Beghi, Capretta, & Soldavini, 2002; Wang, Hurst, Abramovitch, & Franklin, 1994). For high precision servo-control systems, it is necessary to investigate friction in smaller dimensions, leading to complex continuous static or dynamic models on the micro-level which depend on both position and velocity of the actuator (Dupont, Hayward, Armstrong, & Altpeter, 2002; Ge, Lee, & Ren, 2001; Wang, Ge, & Lee, 2001). One way of dealing with the problems of bias forces and friction is to introduce appropriate servo-control techniques to compensate for these complex forces. Linear techniques such as an increased low frequency controller gain are not sufficient to overcome this non-linear effect. It has been shown that the non-linear effects, such as friction hysteresis, affect in particular head position changes for small span seeks. Hence, non-linear methods are necessary for compensation (Gong et al., 2002; Herrmann, Ge, & Guo, 2005).

In practice, it is very hard to obtain an accurate model for these position-dependent and time-varying forces. Hence, adaptive neural network (NN) compensation appears to be most suitable. The advantages of using NN-techniques for compensation of bias forces and friction have been already tested by Ge, Lee, and Harris (1998, p. 172) in closed-loop adaptation for robotic systems and by Huang, Ding, Weerasooriya, and Low (1998) for HDDs (off-line estimation). Most recently, an adaptive NN-controller with on-line adaptation in closed loop has been tested by Herrmann et al. (2005) for a HDD-servo-control system. This controller was developed in a *continuous-time* framework for manipulators with rigid body dynamics (Ge et al., 1998) and then suitably discretized for implementation. The NN-controller structure combining a linear PI(D)-control element with a non-linear adaptive NN-control term has shown to be effective for a VCM-actuator in a HDD. However, in HDDs, it is necessary to directly address inherent sampled-data issues by developing a discrete-time design approach for a friction compensator. As friction is mostly detrimental to the track-following control task, it is more desirable to improve on the overall performance of an existing well-performing, linear discrete track-following controller. This can be done by adding an adaptive NN-control element. For the HDD-control engineer, it is a common approach to add additional compensators for various disturbance rejection aspects (Huang et al., 1998; Jia, Wang, & Wang, 2005; Sri-Jayantha et al., 2001) as it allows to retain a well-performing controller which is further improved.

In continuous time, the augmentation of an NN-controller to a linear controller has been a popular approach in the control community (e.g. Calise, Yang, & Craig, 2002; Campa, Sharma, Calise, & Innocenti, 2000; Hovakimyan, Yang, & Calise, 2002; Steck, Rkhsaz, & Shue, 1996; Yang, Calise, & Craig, 2003). However, as HDD-servo controllers are sampled-data control systems, it is necessary to develop suitable ideas in discrete time. Thus, the approach is to design at first a linear discrete high performance track following controller which achieves a desired bandwidth combined with suitable stability margins (see for instance Herrmann & Guo, 2004). The second step is then to design a discrete NN-controller for bias force and friction compensation. This NN-control element is combined with the discrete linear high performance controller. From a practical point of view, this appears to be feasible as non-linear bias force and friction effects are usually observed in low frequency (Abramovitch, Wang, & Franklin, 1994) which can be compensated by the non-linear NN-element. Thus, a rationale of splitting operation in high and low frequency action is introduced as an important basis of this work. A custom-designed industrial linear track following controller with good high frequency characteristics is continued to be used primarily acting in the high and middle frequency range and the non-linear control element acts as an 'add-on' for rejection of bias and friction forces in low frequency.

The approach of a discrete NN-controller has to remain simple in HDD-servo-systems. For this very reason, the matched uncertainty approach used in Herrmann et al. (2005) appears to be most suitable. Hence, the bias forces and friction are collocated with the actuating torque of the VCM-actuator (see also Oboe et al., 2002; Peng, Chen, Cheng, & Lee, 2005 for an example of a rigorous simulation model with matched disturbances due to friction). These torque disturbances are forces acting likewise the VCM at the pivot of the actuator arm. Furthermore, Herrmann et al. (2005) have shown that pivot friction and flex cable forces can be regarded as unknown *model-dependent* disturbances and are modelled as continuous functions of the VCM-actuator position and velocity. Thus, NN-control methods are suitable to compensate for them. In contrast, windage, vibration or external shocks are not model-dependent disturbances although they can be regarded as torque disturbance in a wider sense. Hence, the add-on NN-control method is regarded as most beneficial for compensating the model-dependent disturbances, friction and flex-cable bias forces, although there might be also some benefit for the other torque disturbances. Ideas from Ge, Lee, Li, and Zhang (2003) are adopted for discrete analysis of this combination of a linear discrete track following controller and an

adaptive discrete NN-controller. Furthermore, principles from Ge et al. (1998, 2001) and Wang et al. (2001) have been essential to this work. Hence, this article presents the necessary theoretical background and simulation/experimental results showing the effectiveness for this suggested NN-approach.

This paper is structured as follows. In Section 2, the nominal linear controller is explained. The adaptive NN-controller component in combination with the nominal linear control system and a relevant stability result for this non-linear discrete-time control system is presented in Section 3. In Section 4, the approach of compensating for a non-linear matched disturbance is examined in terms of bias and friction forces affecting the VCM-actuator in HDDs. Results showing the effectiveness of the NN-compensator for a PID-controlled VCM-simulation model are provided in Section 5. Experimental results are presented and discussed in Section 6, while the paper is concluded in Section 7.

2. The nominal control system with matched non-linearity

Consider a linear discrete single-input single-output system with matched unknown non-linear disturbance:

$$\begin{aligned}\mathbf{x}_p(k+1) &= \mathbf{A}_p\mathbf{x}_p(k) + \mathbf{b}_p u(k) + \mathbf{b}_p f(\Phi(\mathbf{y}(k))),\\ y(k) &= \mathbf{c}_p\mathbf{x}_p(k),\end{aligned} \tag{1}$$

where u is the control input, y the output measurement, $\mathbf{A}_p \in \mathbb{R}^{n\times n}$ the system matrix, $\mathbf{b}_p \in \mathbb{R}^{n\times 1}$ the input distribution matrix, $\mathbf{c}_p \in \mathbb{R}^{1\times n}$ the output distribution matrix and the function, $f(\Phi(\mathbf{y}(k)))$, $f : \mathbb{R}^{\phi} \to \mathbb{R}$ is the unknown disturbance, which depends on a known function $\Phi : \mathbb{R}^{\psi} \to \mathbb{R}^{\phi}$ of a finite vector $\mathbf{y}(k) = [y(k)y(k-1)\cdots y(k-\psi+1)]^{\mathrm{T}}$ representing a finite interval of the time history of y. Note that f enters the system via $\mathbf{b}_p$, the same input channels as the actuating signal u.[3] Hence, f is a disturbance to the actuator signal and it is therefore *matched* to the range of the actuator. The function Φ is continuous in $\mathbb{R}^{\psi}$ and remains bounded in a compact subset of $\mathbb{R}^{\psi}$, the set for the time history of the output measurement, so that

$$\|\mathbf{y}(k)\| < K : \|\Phi(\mathbf{y}(k))\| < L, K, L > 0,$$

where the Euler norm $\|\cdot\|$ is used. For brevity $\Phi_k = \Phi(\mathbf{y}(k))$.

The major problem in this case here is the lack of exact knowledge of f. However, it is assumed for the uncertainty f, that it is a continuous, sufficiently smooth, non-linear function in Φ and remains bounded for any compact subset in $\mathbb{R}^{\phi}$.

Assume $u(k) = u_{NL}(k) + u_L(k)$ and there exists a linear controller

$$\begin{aligned}\mathbf{x}_c(k+1) &= \mathbf{A}_c\mathbf{x}_c(k) + \mathbf{b}_c(y(k) - d),\\ u_L(k) &= \mathbf{c}_c\mathbf{x}_c(k) + d_c(y(k) - d),\end{aligned} \tag{2}$$

exponentially stabilizing the plant of (1) for $u_{NL} = 0, f = 0$ and $d = 0$ while achieving ultimate boundedness for $f \neq 0$ and bounded $d \neq 0$. The exogenous signal d is an output disturbance or a closed-loop demand. The control input term $u_{NL}(k)$ is used later to efficiently compensate for the non-linearity $f_k = f(\Phi(\mathbf{y}(k)))$.

Remark 1. Note that practical systems are usually strictly proper. Hence, it is always reasonable to assume that a plant model is strictly proper by introducing a very fast first order pole achieving the strictly proper characteristic. For the high frequency range, it is then assumed that the model is not reliably determined creating some extra, possibly very minor model uncertainty.

The closed-loop system is represented by

$$\begin{bmatrix}\mathbf{x}_p(k+1)\\ \mathbf{x}_c(k+1)\end{bmatrix} = \underbrace{\begin{bmatrix}\mathbf{A}_p + \mathbf{b}_p d_c \mathbf{c}_p & \mathbf{b}_p\mathbf{c}_c\\ \mathbf{b}_c\mathbf{c}_p & \mathbf{A}_c\end{bmatrix}}_{\mathbf{A}_g}\underbrace{\begin{bmatrix}\mathbf{x}_p(k)\\ \mathbf{x}_c(k)\end{bmatrix}}_{\mathbf{x}_g(k)} + \underbrace{\begin{bmatrix}\mathbf{b}_p\\ 0\end{bmatrix}}_{\mathbf{b}_g}(u_{NL}(k) + f_k) - \begin{bmatrix}\mathbf{b}_p d_c\\ \mathbf{b}_c\end{bmatrix} d, \tag{3}$$

where the measurable output of the overall closed loop is

$$\mathbf{y}_g(k) = \underbrace{\begin{bmatrix}\mathbf{c}_p & 0\\ 0 & I\end{bmatrix}}_{\mathbf{C}_g}\begin{bmatrix}\mathbf{x}_p(k)\\ \mathbf{x}_c(k)\end{bmatrix}, \quad \mathbf{y}(k) = \underbrace{[\mathbf{c}_p \;\; 0]}_{\tilde{\mathbf{C}}_g}\begin{bmatrix}\mathbf{x}_p(k)\\ \mathbf{x}_c(k)\end{bmatrix}.$$

This defines the nominal discrete system $(\mathbf{A}_g, \mathbf{b}_g, \mathbf{C}_g)$ for which a non-linear NN term is designed to compensate for the non-linearity f.

3. The non-linear adaptive compensator

This section has been inspired by the work of Ge et al. (2003) on the discrete NN-control for a special class of non-linear systems. Hence, the approach by Ge et al. (2003) has been simplified and suitably adjusted to suit the closed-loop system $(\mathbf{A}_g, \mathbf{b}_g, \mathbf{C}_g)$ with unknown uncertainty/disturbance f.

Since f is continuous, sufficiently smooth in Φ, the non-linearity $f = f(\Phi)$ can be arbitrarily closely modelled (Ge et al., 1998; Wang et al., 2001) in an arbitrary, compact set of $\Phi \in \mathbb{R}^{\phi}$ via

$$f(\Phi) = \mathbf{w}^{*\mathrm{T}}\mathbf{s}(\Phi) + \varepsilon \tag{4}$$

for some NN-basis function vector $\mathbf{s}(\Phi) \in \mathbb{R}^{l}$ and a, respectively, large enough number $l \in \mathbb{N}^{+}$ of NN-nodes. The vector $\mathbf{w}^*$ is the optimal weight vector while the NN-basis function vector satisfies $\mathbf{s}^{\mathrm{T}}(\Phi)\mathbf{s}(\Phi) \leqslant l$.[4] The scalar

[3] Some more discussion on this will follow later in terms of the HDD problem.

[4] For the bounded neural functions of $\mathbf{s}(\Phi)$, it is common to use radial basis functions, e.g. Gaussian radial basis functions, sigmoid functions or (inverse) Hardy's functions (Ge et al., 1998) to model $f(\Phi)$ in a compact set of Φ. For this compact set, each set of NN-basis functions can be orthonormalized, obtaining a complete, orthonormal set in the Hilbert

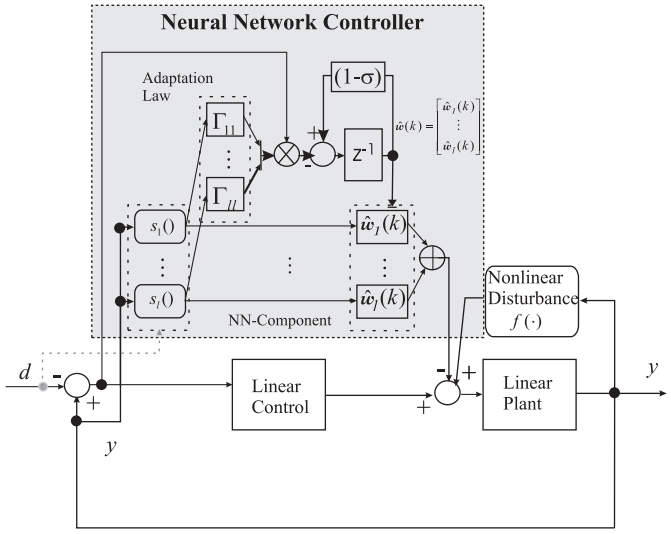

Fig. 1. Linear controller combined with an NN-control component.

$\varepsilon \in \mathbb{R}$ is the modelling error for which $|\varepsilon|$ decreases for an increasing number l of NN-nodes.

Consider the non-linear controller u_{NL} as follows:

$$u_{NL} = -\hat{\mathbf{w}}^{\mathrm{T}}(k)\mathbf{s}(\Phi_k). \tag{5}$$

Hence, the non-linearity is cancelled by an estimate $\hat{f}(k) = \hat{\mathbf{w}}^{\mathrm{T}}(k)\mathbf{s}(\Phi_k)$ for which the weight update law is

$$\hat{\mathbf{w}}(k+1) = (1-\sigma)\hat{\mathbf{w}}(k) - \boldsymbol{\Gamma}\mathbf{s}(\Phi_k)(y(k)-d), \quad \sigma > 0 \tag{6}$$

for a positive definite symmetric matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{l\times l}$, $\boldsymbol{\Gamma} > 0$, and a forgetting factor σ, $0 < \sigma < 1$. Hence, the NN weight estimation error for the weights $\mathbf{w}^*$ is defined by

$$\tilde{\mathbf{w}}(k) = \hat{\mathbf{w}}(k) - \mathbf{w}^*. \tag{7}$$

Using this approach, stability of the overall control system with the NN-controller as depicted in Fig. 1 can be summarized for vanishing output disturbance $d = 0$ by the following theorem:

Theorem 1. *There always exist NN-controller parameters* $1 \gg \sigma > 0$ *and* $\lambda_{max}(\boldsymbol{\Gamma})$ *small enough and* $\lambda_{min}(\boldsymbol{\Gamma}^{-1})\sigma \gg 0$ *large enough so that the linear closed-loop control system from* (3) *in combination with the NN-controller from* (5) *and a weight update law as from* (6) *is ultimately bounded stable.*[5]

The proof of Theorem 1 is provided in the Appendix discussing also the possibility for optimization of the two adaptation law parameters $\boldsymbol{\Gamma}$ and σ. The proof gives insight into the reasons for the choice of parameters which guarantee stability and performance of the closed-loop system.

(*footnote continued*)
space of square integrable functions. However, note that only a maximum of l of these NN-basis functions is used in practice.

[5]Ultimate boundedness is a concept (Khalil, 1992, Definition 5.1) to describe stability in practical systems. In abbreviated form: A system is ultimately bounded stable if any system state trajectory starting in a ball B_α around the origin will enter after finite time a ball B_b around the origin and remain in B_b. (It is often practically desirable that B_b is small.)

Remark 2. Persistent excitation of the NN-estimation algorithm, adapting in closed loop, is not necessary. However, it will be discussed for the experiment later that excitation caused indirectly via disturbances can be beneficial.

Remark 3. The analysis for Theorem 1 includes only matched model-dependent disturbances, while unmatched disturbances or the case of a bounded exogenous signal d have not been considered for reasons of simplicity. However, the introduction of a forgetting factor $\sigma > 0$ (6) has been known to provide robust stability in case of other unaccounted bounded disturbances (Ge et al., 1998). Also in the case here, it can be sufficiently shown for unmatched bounded disturbances that the introduction of the forgetting factor guarantees robustness and ultimate bounded stability. For instance for the exogenous signal d, a bounded tracking error for the control system with NN-control component could be obtained for large $\lambda_{min}(\boldsymbol{\Gamma}^{-1})\sigma$ and small values of $\lambda_{max}(\boldsymbol{\Gamma})$ and σ.

Remark 4. Note that the parameter choice of a large $\lambda_{min}(\boldsymbol{\Gamma}^{-1})\sigma$ and small values for $\lambda_{max}(\boldsymbol{\Gamma})$ and σ guarantees stability but effectively equates to disabling the adaptive NN-controller, achieving nominal linear controller performance only. Hence in this case, the learning algorithm is chosen as slowly acting as possible retaining the NN-weight estimates $\hat{\mathbf{w}}(k)$ close to its initialvalue $\hat{\mathbf{w}}(0)$.

In practice, it is of interest to obtain an NN-controller which acts *improving* on the overall control system behaviour. For this reason, recall the general knowledge for adaptive NN control in continuous-time systems, e.g. Ge et al. (1998). For fast adaptation, it is usually of benefit to obtain large adaptation gains $\boldsymbol{\Gamma}$ and for small ultimate bounded set, it is advisable to have σ small. However, in continuous-time control, it is possible to cancel the modelling error ε (4) via high gain sliding mode methods to achieve asymptotic tracking for a suitable plant output measurement. In a discrete control system, it is not possible to compensate completely for such a constant bounded disturbance ε. A small choice of σ and large values for $\boldsymbol{\Gamma}$ would minimize the set of ultimate boundedness. However, these requirements would equate of having a high gain adaptation algorithm with minimal or no forgetting factor σ, which is possible for continuous-time adaptive NN-control but not for discrete-time control as it would violate respective stability constraints.

Remark 5. The fact that persistent excitation is not necessary for the robust stability of this controller implies that the adaptation algorithm cannot guarantee that the estimation error $\tilde{\mathbf{w}}$ (7) converges to zero (e.g. Ge et al., 2003). This is also a well-known fact for a large class of continuous-time adaptive controllers (e.g. Ge et al., 1998), where only boundedness of the estimation error is guaranteed. In the continuous-time case, it is possible to achieve asymptotic tracking for a suitable plant output measurement despite bounded estimation error. For the

discrete-time case, the control error is guaranteed to be bounded only (see also Remark 4).

Remark 6. In particular for the case of the HDD, the demand d is usually a piecewise constant value. It easily follows that a constant value for d (and also a bounded value of d) will always imply an ultimate bounded stability result. In the practical case, it might be useful to feed a constant demand value d to the NN functions $s(\cdot)$ (see Fig. 1).This has the effect that the operation point of the NN-compensator can be defined to be at the HDD track center given by the demand d. This approach can minimize the time needed for closed-loop adaptation.

In the next section, the concept of the NN-controller will be used to counteract non-linear bias effects on a linear track following controller of a HDD-servo-system.

4. Bias effects in track following of a HDD-servo-system

In HDDs, bias effects are usually caused by bias forces such as forces due to the flex cable or the pivot bearing friction. On a macro-level, friction forces or other bias forces appear to be constant. However, on the micro-level, a friction force requires a much more complex model: Rather complex continuous functions depending on velocity and position of the actuated system are usually employed to model friction characteristics (Dupont et al., 2002; Wang et al., 2001). NN-techniques have shown to be very powerful when modelling these forces in dependence on pivot velocity and position measurement (Herrmann et al., 2005).

For application of the NN-approach from Section 3, it is important that these forces are matched to the range space of the actuator, i.e. the forces enter at the same point as the actuating force of the VCM-actuator. In this case, the non-linear control signal u_{NL} can compensate for them. For the track following loop of a HDD-servo-system, these forces are most of the torque disturbances, such as the pivot non-linearities due to friction or bias forces due to the flex cable (see Fig. 2). These forces can be also reasonably well modelled as functions of VCM-pivot velocity and position (see Oboe et al., 2002; Wang et al., 1994; Peng et al., 2005 for a discussion of bias forces, e.g. friction, acting as matched model-dependent torque disturbances). Furthermore, windage, vibration or external shocks are usually regarded as matched torque disturbances affecting the VCM-system at the pivot of the actuator (see for instance Yang, Jeong, Park, & Park, 2001). Although there might be some potential of counteracting these non-model based torque disturbances (as long as they are not fast), the add-on NN-control method is regarded as most beneficial for compensating the model-dependent matched disturbances, friction and flex-cable bias forces. It is also to mention that it is necessary here to approximate the pivot position and velocity for the disturbance model by the head position error signal and velocity. It can be assumed here that the VCM-actuator is sufficiently stiff to allow the application

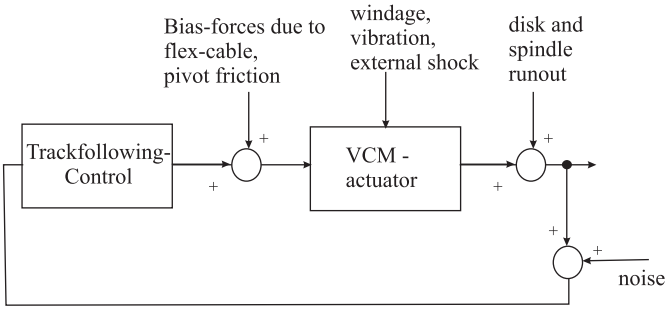

Fig. 2. 'Standard' block diagram of track-following HDD-loop.

of the NN-technique for compensation of friction and other bias forces. This practical assumption has been successfully validated by Herrmann et al. (2005) and Huang et al. (1998) and it shows to be very effective in this paper. Note also that the NN-control algorithm does not compensate for the unmatched disturbances, such as noise and run-out (see Fig. 2), although it is sufficiently robust against these disturbances (see Remark 3).

Considering the discussion above, the model for the bias and friction force $f = f(\Phi)$ in particular for Φ is

$$\Phi_k = \begin{bmatrix} y(k) - d \\ y(k) - y(k-1) \end{bmatrix} \approx \begin{bmatrix} y(t) - d \\ T_s \dot{y}(t) \end{bmatrix},$$

where T_s is the sampling time of the practical sampled-data control system. As for Wang et al. (2001), f may be modelled using Gaussian radial basis functions for the NN-basis functions of $\mathbf{s}(\Phi_k) = [s_1(\Phi_k) s_2(\Phi_k) \cdots s_l(\Phi_k)]^{\mathrm{T}}$:

$$s_i(\Phi_k) = \mathrm{e}^{-(y(k)-d-c_{y/i})^2/\sigma_{y/i}^2} \cdot \mathrm{e}^{-(y(k)-y(k-1)-c_{\Delta y/i})^2/\sigma_{\Delta y/i}^2}. \tag{8}$$

The values of $\sigma_{y/i}, \sigma_{\Delta y/i} \in \mathbb{R}^+$ and $c_{y/i}, c_{\Delta y/i} \in \mathbb{R}$ are the variance and center positions for the position measurement and the velocity estimate of the Gaussian radial basis functions. These parameters need to be chosen by the designer. Note that the value of the demand d is also included in the model of the NN-basis function. This approach takes account of the fact that d is piecewise constant and that the Gaussian radial basis function is locally valid. If the information of d is not given to the NN-basis function, the values of the radial basis functions may become small, $\mathbf{s}(\Phi_k) \approx 0$, for large values, $d \gg \sigma_{\Delta y/i}$, and the NN-control law may be ineffective. Hence, the value of d is included in $\mathbf{s}(\Phi_k)$; fast adaptation will correct the weight for $s_i(\Phi_k)$. The NN is centered around the demanded track.

5. Simulation study for a non-linear HDD-model

To allow a detailed analysis of the NN-adaptation algorithm, a practical implementation and a simulation study is carried out. The advantage of carrying out a simulation study is that the usually unknown bias force disturbance $f(\cdot)$ (see Eq. (1)) is available and can be related to the estimated value $\hat{f}(k) = \hat{\mathbf{w}}^{\mathrm{T}}(k)\mathbf{s}(\Phi_k)$. In this case, a recently established non-linear model for a VCM-actuator (Peng et al., 2005) is used to evaluate the

control algorithm:

$$\ddot{\bar{y}} = 2.35 \times 10^8 u + 2.35 \times 10^8 f_V, \tag{9}$$

where the bias force f_V is

$$f_V = -0.02887 \arctan(0.5886\bar{y}) + \tilde{T}_f$$

and

$$\tilde{T}_f = 4.255 \times 10^{-9} \begin{cases} -[|1.175 \times 10^6 \bar{y} u + 0.01(\dot{\bar{y}})^2| + 15\,000] \operatorname{sign}(\dot{\bar{y}}) & \\ -282.6\dot{\bar{y}} & \dot{\bar{y}} \neq 0, \\ -\tilde{T}_e & \dot{\bar{y}} = 0, \ |\tilde{T}_e| \leqslant \tilde{T}_s, \\ -\tilde{T}_s \operatorname{sign}(\tilde{T}_e) & \dot{\bar{y}} = 0, \ |\tilde{T}_e| > \tilde{T}_s, \end{cases}$$

$$\tilde{T}_e = 2.35 \times 10^8 [-0.02887 \arctan(0.5886\bar{y}) + u], \tag{10}$$

$$\tilde{T}_s = 1.293 \times 10^6 |u_0 \bar{y}_0| + 1.65 \times 10^4. \tag{11}$$

The values $\bar{y}_0$ and u_0 are the corresponding input and displacement for $\dot{\bar{y}} = 0$. This establishes the dynamics of the pivot $G_p : u \mapsto \bar{y}$, while the actuator flexibilities are

$$\begin{aligned} G_r = {} & \frac{0.8709s^2 + 1726s + 1.369 \times 10^9}{s^2 + 1480s + 1.369 \times 10^9} \\ & \times \frac{0.9332s^2 - 805.8s + 1.739 \times 10^9}{s^2 + 125.1s + 1.739 \times 10^9} \\ & \times \frac{1.072s^2 + 925.1s + 1.997 \times 10^9}{s^2 + 536.2s + 1.997 \times 10^9} \\ & \times \frac{0.9594s^2 + 98.22s + 2.514 \times 10^9}{s^2 + 1805s + 2.514 \times 10^9} \\ & \times \frac{7.877 \times 10^9}{s^2 + 6212s + 7.877 \times 10^9}. \end{aligned} \tag{12}$$

The overall plant model G_{vcm} is given by the composition, $G_{vcm} = G_r \circ G_p$. In Peng et al. (2005), two controllers are discussed, a non-linear control approach and a PID-controller, which is enhanced by a non-linear compensator. The PID-controller (Peng et al., 2005), but without the compensator of Peng et al. (2005), shall be the basis for this study:

$$u_L = -\left(k_p + \frac{k_d(z-1)}{T_s\left(\frac{z}{N} + z - \frac{1}{N}\right)} + \frac{k_i T_s z}{z-1} \right)(y - d), \tag{13}$$

where $k_p = 0.0605$, $k_d = 2.8876 \times 10^{-5}$, $k_i = 50$ and $N = 10$. The controller sampling time is $T_s = 1/10\,000$ s. The adaptive NN-controller is now applied to counteract the effect of the matched disturbance of f_V. Friction effects of this HDD model are in particular significant in the sub-micrometer range.

5.1. NN-controller design

It is to be proven that the adaptive NN-algorithm is effective for a small number of NN-nodes. Thus, a number of only three nodes are chosen. Note that it is not necessarily important to fine-tune the characteristics of each NN-basis function. It is more important to distribute the NN-nodes over the controller's operating region. Practical tests can reveal the best choice of NN-nodes.

Friction compensation has to be effective in the sub-micrometer range. Hence, the centers $c_{y/i}$, $c_{\Delta y/i}$ of the radial basis functions are chosen in this range:

$$\begin{aligned} & c_{y/1} = -0.167\,\mu\text{m}, \quad c_{y/2} = 0.167\,\mu\text{m}, \quad c_{y/3} = 0\,\mu\text{m}, \\ & c_{\Delta y/1} = -0.033\,\mu\text{m}, \quad c_{\Delta y/2} = 0.033\,\mu\text{m}, \quad c_{\Delta y/3} = 0\,\mu\text{m}. \end{aligned} \tag{14}$$

The node, no. 3 ($c_{y/3} = 0$, $c_{\Delta y/3} = 0$), is the 'central' node which is used, while the other two nodes should be representative of the off-center characteristics of the system. The value range of the centers $c_{\Delta y/i}$ can be found by investigating the difference $y(k) - y(k-1)$ in simulation tests. In particular, the nominal controller has been investigated for step responses in the range of 100 nm. The amplitudes of $y(k) - y(k-1)$ are in the range of 100 nm, so that the centers $c_{\Delta y/i}$ are chosen accordingly. The radial basis function have to sufficiently overlap each others' range so that $\sigma_{y/1} = \sigma_{y/2} = \sigma_{y/3} = 0.158\,\mu$m, and $\sigma_{\Delta y/1} = \sigma_{\Delta y/2} = \sigma_{\Delta y/3} = 0.0707\,\mu$m. The forgetting factor $\sigma = 0.01 \cdot T_s$ is chosen in this case. This is a very small value and admissible for this simulation. However, in a practical application, it is generally better that the value of σ is larger to retain robust stability.

5.2. Simulation results and discussion

To show the feasibility of the approach, the small step responses are assessed in the range up to 1 μm (see Figs. 3 and 4). From Figs. 3(a) and 4(a), it can be seen that the nominal controller tested for $f_V = 0$ (9) has been designed with a settling time of about 2 ms and an overshoot of about 0.6. Hence, any non-linear compensator achieving these characteristics has recovered the nominal controller in a reasonable manner. This is for instance the case in Peng et al. (2005). In Figs. 3(a) and 4(a), it can be seen that the fully uncompensated PID-controller is subject to a long settling time of more than 3 ms. The non-linear NN-compensation scheme is generally well able to compensate for this loss of performance due to the non-linear effects of the VCM-actuator model. Hence, settling times of about 2 ms are recovered and the overshoot is generally smaller than 0.6. Note that the NN-basis function weights are initialized with 0. Hence, adaptation is fast, although gradual tracking improvement can be observed over a longer time period (see Fig. 4(a)).

Note also that the NN-adaptation does not necessarily guarantee convergence to the disturbance f (see Eq. (1), see

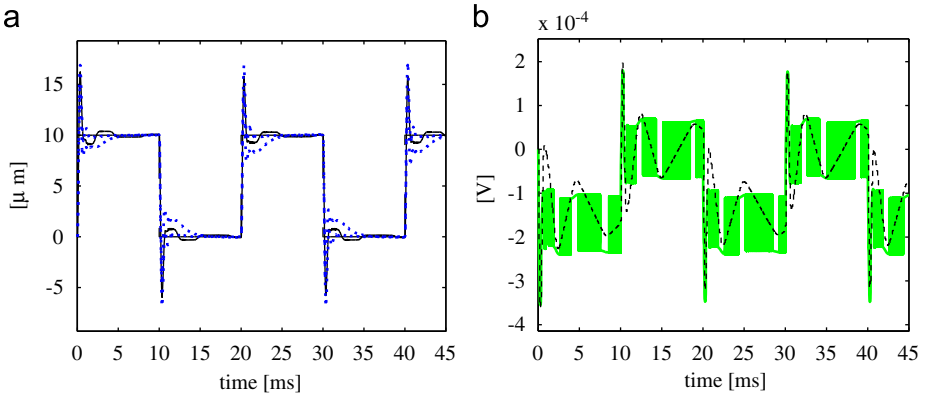

Fig. 3. Step response of simulation model under PID-control (Peng et al., 2005) with/without NN-compensation or non-linearity: (a) Output measurement; dashed, demand; line, NN-compensated; dotted, not NN-compensated; dot-dashed, nominal control without NN-compensation and for $f_V = 0$. (b) Disturbance f_V (light line) and estimate $\hat{f}$ (dashed dark) for NN-compensated control.

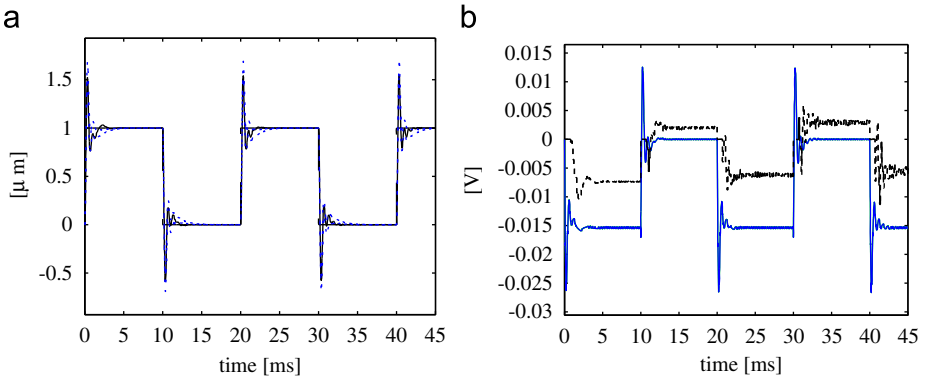

Fig. 4. Step response of simulation model under PID-control (Peng et al., 2005) with/without NN-compensation or non-linearity: (a) Output measurement; dashed, demand; line, NN-compensated; dotted, not NN-compensated; dot-dashed, nominal control without NN-compensation and for $f_V = 0$. (b) Disturbance f_V (line) and estimate $\hat{f}$ (dashed) for NN-compensated control.

Remark 5). However, it is here for small step demand obvious that reasonable convergence can be achieved (Fig. 3(b)) while the estimation error for large step remains bounded only (Fig. 4(b)).

5.3. Discussion

It is seen that the adaptive NN-compensator can recover nominal controller performance over a large range of small signal step responses.[6] This has been also achieved for the PID-control by Peng et al. (2005, Fig. 12) where the compensator for the PID-control is model based. This shows that the NN-compensator, which requires very little information about the practical system, is as effective.

Practical studies in the next section will show that this is also the case in a realistic HDD-servo experiment.

[6]Note that significantly larger step responses are usually carried out with a suitable non-linear seek method which allow the scheduling of a time-optimal seek-settling controller and smoothly changes over into a linear track-following controller.

6. Practical implementation of the NN-controller for friction and bias-force compensation

The experiments have been conducted with an up-to-date high performance 3.5 in VCM-actuator[7] to which a 23 mm suspension[8] with read–write head was attached. For the practical tests, an A*Star DSI designed spin stand was used to mount the read–write head on a commercial 2.5 inch disk rotating at 3300 RPM. The (horizontal) position of the read–write head is measured with a Laser-Doppler-Scanning-Vibrometer[9] (LDV) (Fig. 5). This creates a highly realistic environment for HDD-experiments so that the VCM-actuator is subjected to the most significant disturbances usually observed in HDD-systems (see Fig. 2), such as disk and spindle runout, windage and bias forces.[10]

[7]Seagate Technology.
[8]A*Star DSI.
[9]Polytec OFV 3001S, Polytec, Waldbronn, Germany.
[10]Demodulation noise is not observed as the VCM-position is recorded with an LDV.

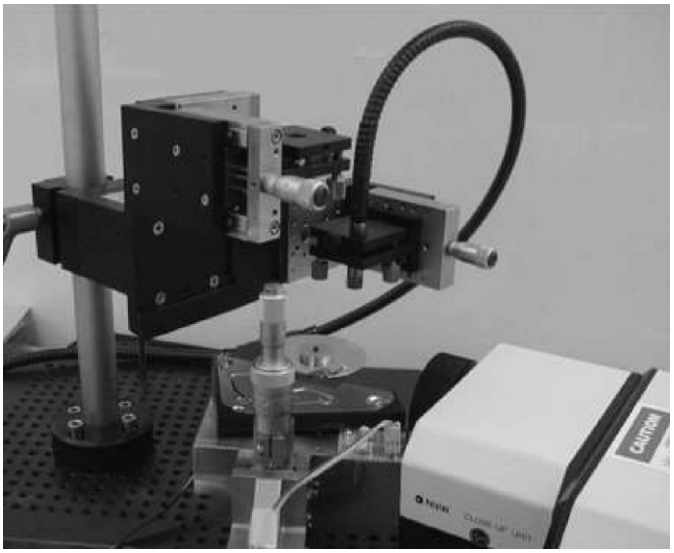

Fig. 5. Experimental set-up.

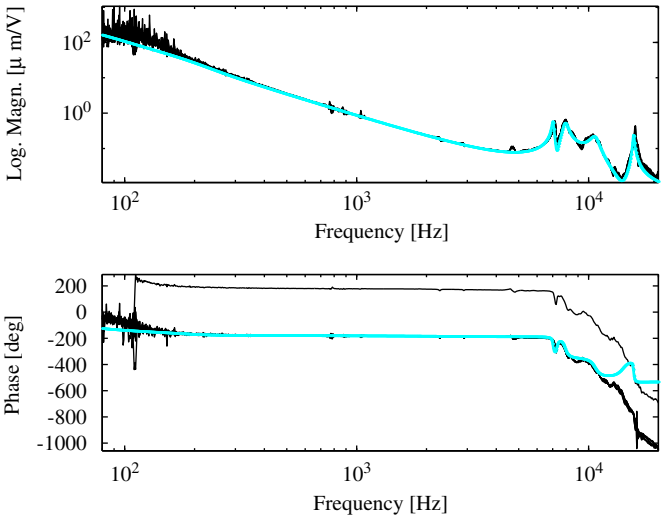

Fig. 6. VCM-frequency responses (including voltage-to-current driver); 10 measurements for five different amplitudes: dark; nominal 12th order transfer function: light.

Employing a Dynamic Signal Analyser[11] (DSA) and five different excitation amplitudes for 10 measurements, the VCM-model measurements show in Fig. 6 below 100 Hz the typical variation in the frequency response amplitude caused by non-linear effects, such as friction and other bias forces. From the frequency response measurements for the actuator, we derived from the continuous measurement the linear continuous nominal VCM-actuator models of 12th order via curve fitting (Fig. 6). Using robust linear μ-design techniques (see Herrmann & Guo, 2004), this allowed to design a discrete linear track following controller of 10th order:

$$\frac{5.52(z-0.88)(z+1)(z^2+1.99z+0.9978)(z^2+1.91z+0.94)(z^2-0.049z+0.94)(z^2-0.45z+1.03)}{(z^2+1.83z+0.84)(z^2+1.97z+0.98)(z^2+1.55z+0.7932)(z^2-0.75z+0.37)(z^2-0.8z+0.84)}$$

with an open-loop crossover frequency at 1460 Hz and a phase and gain margin of 33° and 4.7 dB (Fig. 7). The controller has been implemented with the DSP based system, DS1103,[12] employing a sampling frequency of $1/T_s = 33$ kHz. Using this nominal linear controller configuration, the NN-controller has been tuned. A description for this follows in the next sections.

6.1. NN-controller design

The characteristics of each NN-node are provided in Table 1. Assuming good knowledge about the practical system, i.e. about the amplitude range of the position and velocity signal during closed-loop control, a good preliminary NN-controller is easily obtained. It will be seen that the number of required NN-nodes for bias and friction force compensation can be retained small for the investigated case, minimizing the effort for on-line tuning of the NN-controller.

The first step for design is to choose an NN-node with centers $c_{y/i}$ and $c_{\Delta y/i}$ directly at the operation point, $c_{y/i} = 0$, $c_{\Delta y/i} = 0$, for which the NN-node $i = 5$ has been selected in Table 1. This node at the operation point is the 'central' node which is used for all investigated NN-configurations. As the next step, four nodes where selected which are symmetrically distributed in close proximity of the operation point, i.e. NN-nodes $i = 2, 3, 7, 8$. To allow also a good and accurate model slightly more further off from the operation point, four other NN-nodes $i = 1, 4, 6, 9$ are to be tested. From these nine nodes, NN-node-configurations are selected which are always symmetric with respect to the central node. The variance values are $\sigma_{y/i} = \frac{20}{\sqrt{17}}$ and $\sigma_{\Delta y/i} = \frac{20}{\sqrt{5 \cdot 17}}$ for all nodes to retain a reasonable overlapping of all nine Gaussian radial basis functions.

It has been discussed in Remark 4, that for performance and maximal learning speed it is desirable to choose the learning coefficients $\mathbf{\Gamma} = diag(\Gamma_{1,1}, \Gamma_{2,2}, \Gamma_{3,3}, \ldots, \Gamma_{9,9})$ as large as possible. However, stability constraints in this discrete-time system do not permit these gains to be too large. Thus, the learning coefficients have been chosen on-line, not excessively large to guarantee robust stability (Table 1). Note that the nodes $i = 2, 3, 7, 8$ have the same learning coefficients due to their symmetric placement of the relevant centers $c_{y/i}$ and $c_{\Delta y/i}$. The same holds for $i = 1, 4, 6, 9$.

From Remark 4, it is desirable for good performance to choose the forgetting factor σ small. However, robust stability requirements make it necessary that $\sigma = 4 \cdot T_s$. The value has been tuned on-line. Note that a scaling

[11] HP 35670A, Hewlett Packard Company, Washington.

[12] DSpace DS1103 is a product of dSPACE GmbH, Paderborn, Germany.

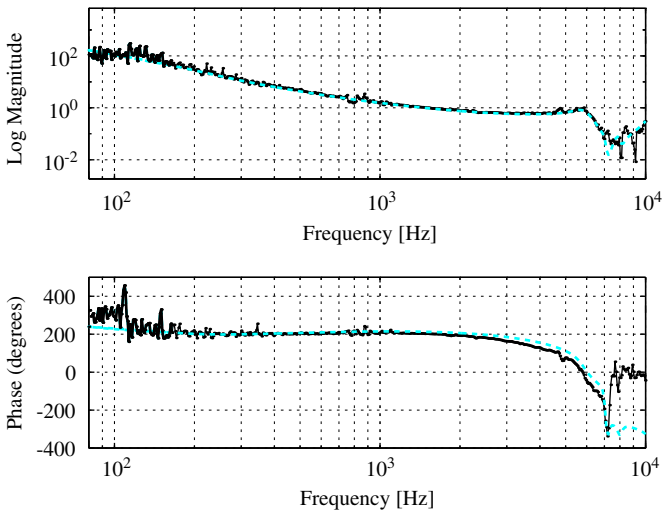

Fig. 7. Open-loop responses—nominal design: light dashed; measured: thick black dot-dash.

Table 1
NN-node characteristics in dependence on the Node No. i ($T_s = 1/33\,000$)

i	$c_{y/i}$ (μm)	$c_{\Delta y/i}$ (μm)	$\Gamma_{i,i}$
1	−1	−0.2	$35{\cdot}T_s$
2	−0.333	−0.0666	$60{\cdot}T_s$
3	0.333	0.0666	$60{\cdot}T_s$
4	1	0.2	$35{\cdot}T_s$
5	0	0	$200{\cdot}T_s$
6	−1	0.2	$35{\cdot}T_s$
7	−0.333	0.0666	$60{\cdot}T_s$
8	0.333	−0.0666	$60{\cdot}T_s$
9	1	−0.2	$35{\cdot}T_s$

Table 2
NN-controller combinations

No. of nodes	1	3	3	9
Active Node No.	5	2,3,5	1,4,5	$1,2,\ldots,9$

factor for the learning coefficient and the forgetting factor is the sampling time $T_s = 1/33\,000$ s.

Overall, it has been found that adaptation is achieved within 1–3 ms. Note that, the forgetting factor $\sigma = 4 \cdot T_s = 0.12$ ms is small, retaining the influence of this parameter small in relation to the adaptation speed and process.

Using these design characteristics, several controller combinations with one NN-node, three NN-nodes or nine NN-nodes have been tested. For these controllers the results are provided in the next section (see Table 2).

6.2. Experimental results

Three different tests have been conducted:

(1) Small step responses.
(2) Error rejection measurement through sine-sweep excitation of the exogenous signal d using a DSA.

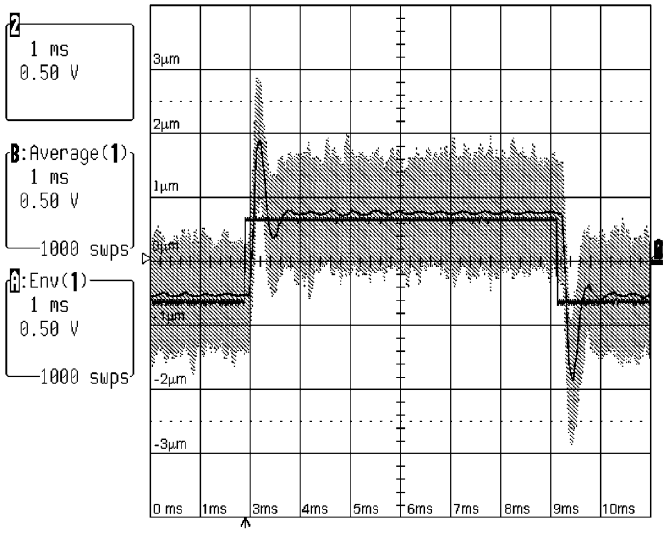

Fig. 8. Step response for linear controller (2 μm/V) (response is averaged; includes envelope of 1000 responses (grey shaded)).

(3) FFT measurement of the position error for $d = 0$ using a DSA.

All of the three different *experiments confirm* that the *NN-controller can improve track following* of the original linear track-following scheme, in particular, error rejection in the low frequency region is increased.

6.2.1. Small step responses

A comprehensive study for small step responses of an amplitude of about 1 μm considering 1, 3 and 9 NN-nodes for the adaptive compensator is carried out.[13] The VCM-actuator is set up in a realistic environment lacking only demodulation noise as disturbances. These observed disturbances (windage, disk vibrations, disk flutter, spindle runout, etc.) occupy a wide frequency spectrum (Yang et al., 2001). Thus, it has been decided to record more than 1000 step responses to obtain the average of the responses (see Figs. 8–10). Note also the respective (grey-shaded) envelope which is due to the recorded 1000 step responses. This is observable for both the controller results with and without NN-compensation. The response of interest is the average response as it is free of non-repeatable disturbances.

The result of this experiment is an *improved low frequency error rejection*. Thus, *steady state errors* are significantly *reduced* using an NN-controller. This can be observed for all the step responses (see Fig. 8 versus Figs. 9 and 10), once the NN-control element is activated.

6.2.2. Tracking error for small amplitude, sinusoidal demands—error rejection

Error rejection measurements are conducted using demand excitation amplitudes of 200 nm for the sine sweep

[13] See footnote 6.

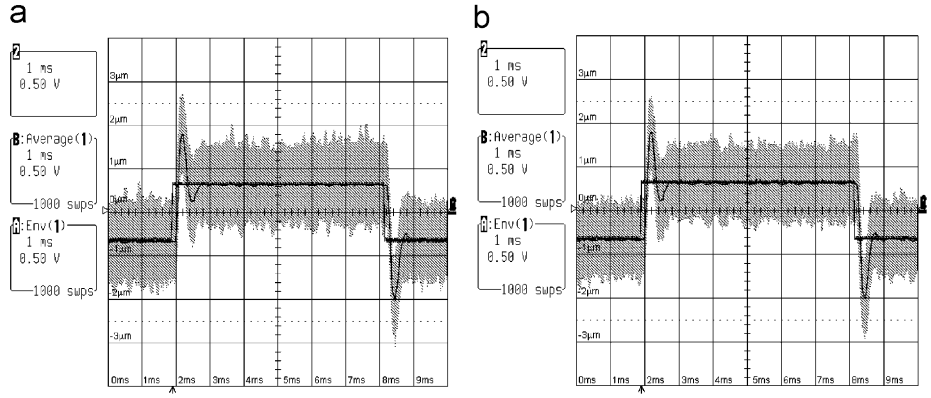

Fig. 9. Step response (2 µm/V) with decreased steady state error in relation to linear control of Fig. 8 (response is averaged; includes envelope of 1000 responses (grey shaded)): (a) three NN-nodes (Node No. 2,3,5); (b) three NN-nodes (Node No. 1,4,5).

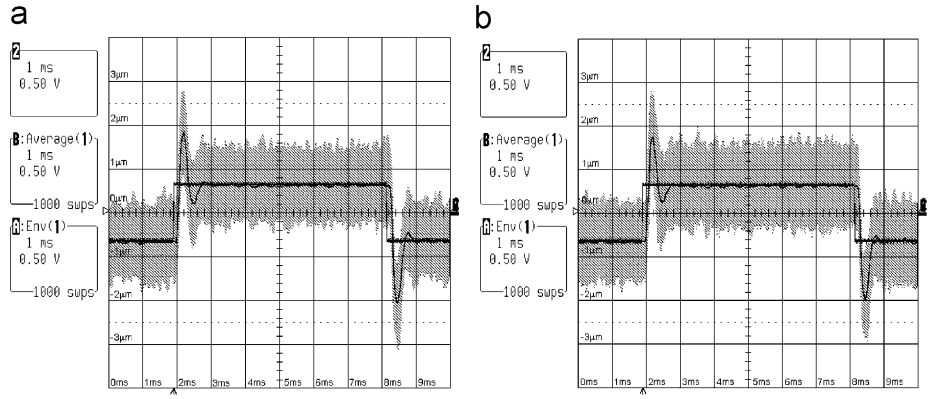

Fig. 10. Step response (2 µm/V) with decreased steady state error in relation to linear control of Fig. 8 (response is averaged; includes envelope of 1000 responses (grey shaded)): (a) one NN-node (Node No. 5); (b) nine NN-nodes (Node No. 1, ..., 9).

signal of the DSA. Note that the term of error rejection has been used here rather than sensitivity response. The system investigated here represents a non-linear closed-loop system while the term sensitivity response (but also error rejection) is usually attributed to linear closed-loop systems, where the frequency response of a single degree of freedom controller from demand to error is investigated. In this paper, the excitation amplitude is of significance as the control system is non-linear; the term error rejection appears to suit the non-linear domain better. Small demand excitation amplitudes have been used in the range of the track-width of a 100 kTPI drive; this is also the operating range for which the NN-compensator has been designed. This allows a very realistic evaluation of the capabilities of the control scheme in the demand amplitude range which is significant for track-following performance and small seeks, i.e. the HDD-servo tasks which usually require a significant understanding of the error rejection response. Such a test has been also conducted in Herrmann et al. (2005) to exemplify the capabilities of a non-linear adaptive control scheme and can be regarded as a practical describing function test (Khalil, 1992) for small amplitudes commonly used in non-linear systems as a generalization of the concept of frequency responses (see Abramovitch et al., 1994; Gong et al., 2002 for HDD related work or Nuij, Steinbuch, & Bosgra, 2008 for more advanced methods). Thus, a limited abuse of usual linear control terminology is necessary for description of the following practical observations.

Improvements in the low frequency region: The NN-controller component improves on the error rejection of the nominal linear controller in particular in the frequency range below 100 Hz, as seen for Fig. 11(b) versus (The amplitudes are the result of a root-mean-square averaging process.) Figs. 12(b), 13(b), 14(b) and 15(b). Improvements for the error rejection of more than 15 dB in the low frequency range are possible. For the NN-controller with nine NN-nodes, almost 20 dB improvement are achieved for the lowest measured frequency at 10 Hz (see Fig. 15(b) versus 11(b)). Note that the NN-controller with three NN-nodes (Node No. 2,3,5) can achieve nearly the same result.

Insignificant changes in the high frequency region: In the high frequency region above 600 Hz, the use of an NN-controller may slightly increase error rejection. However, the effect of the NN-controller remains minimal (see Table 3) in this frequency region. The effects are in particular for the tested NN-controllers with one and three NN-nodes small. The maximal error rejection peak does not increase more than 3.7%, which is true for the controller with three NN-nodes (Node No. 1,4,5). It is even the case, that the maximal error rejection peak

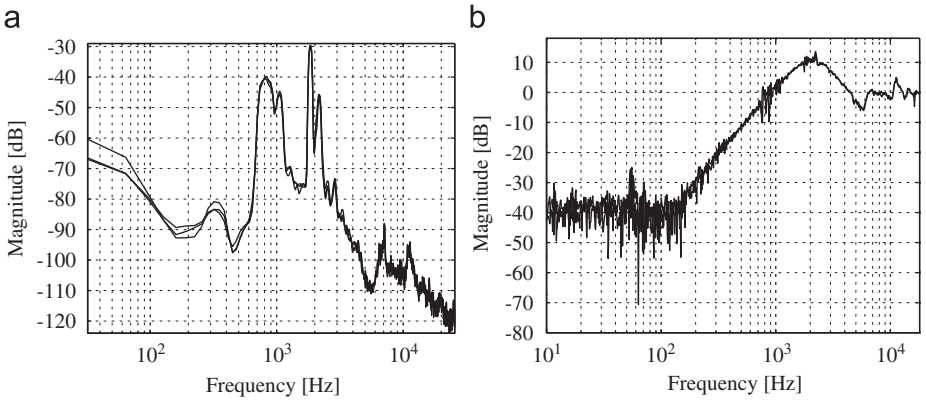

Fig. 11. FFT and error rejection for linear controller only: (a) $|FFT|$ of posit. meas. (μm); (b) error rejection.

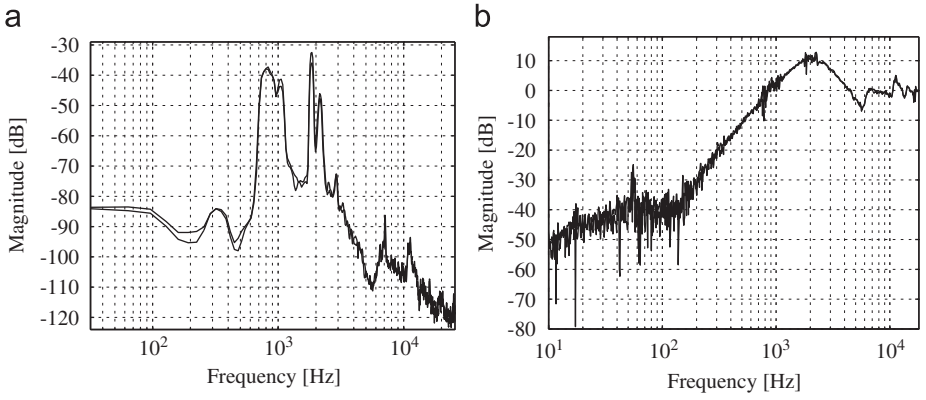

Fig. 12. FFT and error rejection for control with one NN-node (Node No. 5): (a) $|FFT|$ of posit. meas. (μm); (b) error rejection.

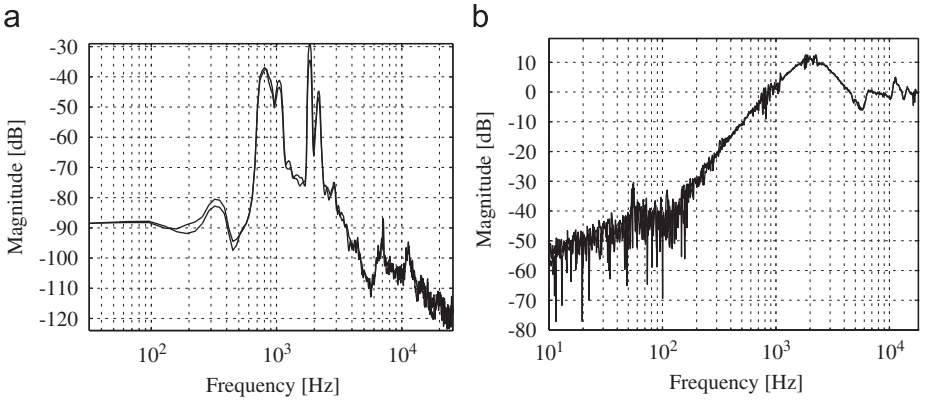

Fig. 13. FFT and error rejection for control with three NN-nodes (Node No. 2,3,5): (a) $|FFT|$ of posit. meas. (μm); (b) error rejection.

decreases for the other NN-controller with three NN-nodes (Node No. 2,3,5) in comparison to the nominal linear controller. A slightly larger increase in the maximal peak is noticed for the NN-controller with nine NN-nodes.

Insignificant changes in the error rejection bandwidth: The error rejection bandwidth[14] remains almost constant for the controllers with one and three NN-nodes when compared to the nominal controller (see Table 3).

Considering the analysis above, the high frequency performance of the nominal linear controller remains almost unaffected by the NN-control. Furthermore, it appears to be *most efficient* to use the *NN-controller* with *three NN-nodes* (Node No. 2,3,5). This can be also verified when observing the FFT-responses.

[14]The bandwidth is determined via the smallest frequency for which the error rejection has a −3 dB value.

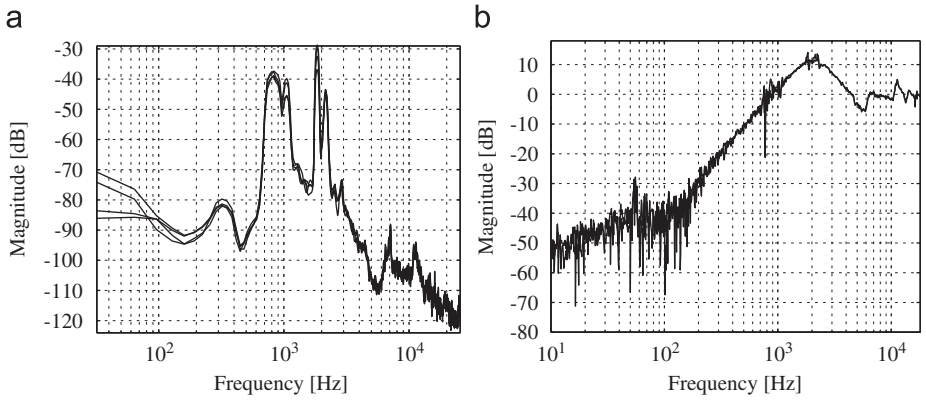

Fig. 14. FFT and error rejection for control with three NN-nodes (Node No. 1,4,5): (a) $|FFT|$ of posit. meas. (μm); (b) error rejection.

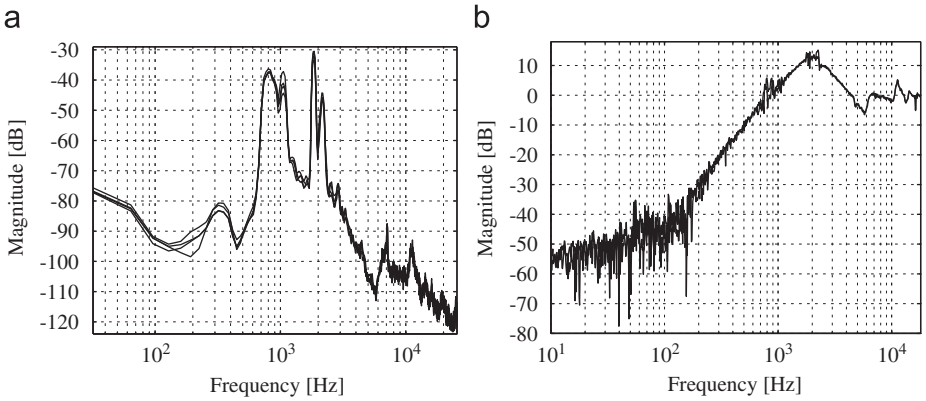

Fig. 15. FFT and error rejection for control with nine NN-nodes: (a) $|FFT|$ of posit. meas. (μm); (b) error rejection.

Table 3
High frequency characteristics in dependence on the NN-characteristics

No. of nodes	Active Node No.	Error rejection		FFT peak (dB)	
		Peak (dB)	Bandwidth (−3 dB) (Hz)	at 830 Hz	at 1850 Hz
0	NA	13.6	713.6	−39.8	−29.6
1	5	12.8	702.6	−37.3	−32.5
3	2,3,5	12.5	703.2	−36.9	−29.1
3	1,4,5	14.1	700.1	−37.3	−28.7
9	1, 2, . . . , 9	15.2	674.0	−36.3	−30.5

6.2.3. FFT measurement

Improvements in the low frequency region: Consider Fig. 11(a) versus Figs. 12(a), 13(a), 14(a) and 15(a) in the frequency range up to 100 Hz. The FFT-amplitude is reduced by more than 10 dB when using an NN-control element. The appropriate choice of NN-nodes assures an even better result of a 20 dB reduction as seen for Fig. 11(a) versus 13(a).

Although for the low frequency error rejection responses in the sub-section before, the error is always reduced for increasing number l of NN-nodes, it is observed for the FFT-recordings that the improvement is much stronger dependent on the choice of the NN-node characteristics. It is important to obtain at first a good model of the non-linear disturbance by selecting the correct NN-node center values close to the operation point (see Figs. 12(a) and 13(a)) while center points $c_{y/i}$ and $c_{\Delta y/i}$ too far off from the operation point can even act slightly destructive (Fig. 13(a) versus Figs. 14(a) and 15(a)). Nevertheless, the NN-controller has been always acting in a *positive and improving manner* in the low frequency region. Furthermore, the experiments showed that *an increase* of the NN-node number l *does not necessarily further improve* low frequency *performance* as it is obvious from Fig. 15 versus Figs. 12–14.

Insignificant changes in the high frequency region: In the high frequency region, the effect of the NN-controller on

the FFT-measurements is also minimal, similar to what was observed for the error rejection responses (see Table 3).

In summary, among those tested configurations, it appears that the NN-control element with three NN-nodes (Node No. 2,3,5) is again the most suitable. The low frequency FFT-amplitude improves by more than 20 dB showing the best performance.

6.3. Discussion

The result appears to indicate that an *NN-model of low complexity* is sufficient to achieve good disturbance rejection. It is in particular of interest to see that an NN-control element with the maximum of nine NN-nodes is not able to improve significantly on the result for three NN-nodes (Node No. 2,3,5). The NN-controller is particularly improving performance in low frequency which is associated with the effects of bias forces and friction (see Table 4).

For a large number of NN-nodes, the results seem to point to the problem of *over-learning* or *over-fitting*: For a large number of NN-nodes, the adaptation algorithm attempts to match the given measurements as accurately as possible creating a highly complex model. In case of noise, this can lead to the modelling of a wrong representation which does not represent the actual physical disturbance phenomenon. Nevertheless, it has been shown that it is possible to choose the right combination of nodes from a set of nodes, well distributed over the operation area. The overall number of available nodes does not need to be large so that it is easily possible to find the 'optimal' choice.

It is also important to note that the *high frequency performance* of the nominal linear controller remains *almost unaffected*. In contrast, a linear integrator element usually used to compensate for low frequency disturbances will always cause a phase loss in the open-loop response of a linear control system and will increase the magnitude of the error rejection frequency response. This has not been found for the non-linear add-on NN-control scheme. Thus, the NN-controller is indeed a suitable augmentation to the existing linear track-following controller, as it was discussed in Section 1.

Adaptation can be achieved within 2–3 ms without prior initialization of the NN weights. This adaptation speed is even higher once the NN weights are close to their optimum reducing to a settling time of less than 1 ms. This has been confirmed in simulation tests. Hence, the add-on NN-algorithm is suitable for track following and short span seek operation, where the optimal value of the NN weights does not change too much.

Table 4
Observed low frequency improvements for the error rejection and the FFT-measurements

Observed improvement	Error rejection at 10 Hz (dB)	FFT at 30 Hz (dB)
Improvement with active nodes (2,3,5)	15	20
Maximal improvement	20	20

As discussed for Theorem 1, persistent excitation of the NN-algorithm is not necessary. However, as for many practically applied adaptive algorithms, it has been noted that external excitation is advantageous. Thus, it has been observed that non-model based disturbances (in particular non-repeatable run-out disturbances), such as windage, noise, disk flutter, bearing defects, etc., appear to improve indirectly the performance of the NN-control element since these disturbances keep the learning process of the NN-controller persistently excited and keep the bias and friction force identification running. This also is based on the fact that the NN-algorithm is robust to these disturbances (Remark 3), so that they can be indeed beneficial as persistent excitation to the presented NN-control scheme, continuously learning in closed loop. However, it is certainly not desirable to have these disturbances too large.

Note that the implementation of the controller has been conducted under conditions which are very close to those in a practical HDD (see beginning of Section 6), i.e. the read/write head was mounted on a rotating hard disk and subjected to the most significant disturbances known in a HDD. Thus, the implementation should be readily transferrable to commercial products, once a more detailed analysis also for other HDDs is conducted. Successful transferability is also assured in particular due to the fact that three NN-nodes are sufficient to obtain a satisfying result and computational effort is therefore retained low.

7. Conclusion

An NN-control strategy has been presented which works as add-on to an existing discrete linear controller. A *theoretical* result has been provided proving stability and performance not worse than the nominal linear controller.

It has been shown in simulation and *practically* for all investigated configurations and any investigated number of NN-nodes, that the *NN-control* technique *improves performance* as it is effective in compensating for low frequency bias effects. These bias and friction forces have been identified as pivot friction, flex cable forces and also to a certain extent as windage or head-disk interaction. As the NN-control element has particularly little effect on the high frequency performance, the NN-controller can be regarded as a true add-on controller to the linear, high performance, track following controller. The reasonably short adaptation time of less than 3 ms is also noteworthy. In particular for good initialization, it has been found that adaptation speed can reduce to values less than 1 ms. Furthermore, the NN-controller appears to be especially effective for the low number of three NN-nodes. It has to be investigated if this in general the case for HDDs, as this can simplify the tuning of the adaptation algorithm considering a small

number of NN-nodes only. If this is guaranteed then transferability to real HDDs should be possible.

Appendix A. Proof of Theorem 1

Using ideas of Ge et al. (2003, Theorems 1 and 2), the proof of Theorem 1 is provided: Consider a symmetric positive definite Lyapunov matrix, $\mathbf{P}>0$, for a Lyapunov function candidate $V(k)=\mathbf{x}_g^{\mathrm{T}}(k)\mathbf{P}\mathbf{x}_g(k)+\tilde{\mathbf{w}}^{\mathrm{T}}(k)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{w}}(k)$ where $\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g-\mathbf{P}<0$. The first step of the proof concentrates on the analysis of the Lyapunov function forward difference $\Delta V(k)=V(k+1)-V(k)$ to derive an expression which allows to show ultimate boundedness. This result has proved to be subject to the negative definiteness of a matrix which is non-linear in the controller parameters. For this non-linear matrix inequality, the second step of the proof provides conditions which show that for a certain choice of parameters the non-linear matrix inequality is always satisfied.

Step 1: Employing (3), (6) and the relation

$$2\tilde{\mathbf{w}}(k)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\hat{\mathbf{w}}(k)=\tilde{\mathbf{w}}(k)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{w}}(k)+\hat{\mathbf{w}}(k)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\hat{\mathbf{w}}(k)-\mathbf{w}^{*\mathrm{T}}\boldsymbol{\Gamma}^{-1}\mathbf{w}^{*},\tag{15}$$

the value of $\Delta V(k)=V(k+1)-V(k)$ can be computed for $d=0$ as

$$\begin{aligned}\Delta V(k)=&\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}^{\mathrm{T}}\mathbf{M}_1\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}+\sigma\mathbf{w}^{*\mathrm{T}}\boldsymbol{\Gamma}^{-1}\mathbf{w}^{*}\\&+2\mathbf{x}_g^{\mathrm{T}}(k)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\varepsilon-2\varepsilon\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\mathbf{s}(\Phi_k)^{\mathrm{T}}\tilde{\mathbf{w}}(k)\\&+\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\varepsilon^2,\end{aligned}\tag{16}$$

where the symmetric matrix $\mathbf{M}_1$ satisfies:

$$\begin{aligned}\mathbf{M}_1=&\begin{bmatrix}\mathbf{M}_{11}&-(\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g+\tilde{\mathbf{C}}_g^{\mathrm{T}})\mathbf{s}^{\mathrm{T}}(\Phi_k)&\tilde{\mathbf{C}}_g^{\mathrm{T}}\sigma\mathbf{s}(\Phi_k)^{\mathrm{T}}\\ *&\mathbf{M}_{12}&0\\ *&*&(\sigma^2-\sigma)\boldsymbol{\Gamma}^{-1}\end{bmatrix},\\ \mathbf{M}_{11}=&\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g-\mathbf{P}+\tilde{\mathbf{C}}_g^{\mathrm{T}}\mathbf{s}^{\mathrm{T}}(\Phi_k)\boldsymbol{\Gamma}\mathbf{s}(\Phi_k)\tilde{\mathbf{C}}_g,\\ \mathbf{M}_{12}=&\mathbf{s}(\Phi_k)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\mathbf{s}(\Phi_k)^{\mathrm{T}}-\sigma\boldsymbol{\Gamma}^{-1}.\end{aligned}\tag{17}$$

It is well known that $\pm2\mathbf{a}^{\mathrm{T}}\mathbf{b}\leqslant\mathbf{a}^{\mathrm{T}}\mathbf{a}+\mathbf{b}^{\mathrm{T}}\mathbf{b}$ for two vectors $\mathbf{a},\mathbf{b}\in\mathbb{R}^h$, $(h\in\mathbb{N}^+)$. Hence, from this inequality and (16), it follows:

$$\begin{aligned}\Delta V(k)\leqslant&\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}^{\mathrm{T}}\mathbf{M}_1\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}+\sigma\mathbf{w}^{*\mathrm{T}}\boldsymbol{\Gamma}^{-1}\mathbf{w}^{*}\\&+\delta_1\mathbf{x}_g^{\mathrm{T}}(k)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g\mathbf{x}_g(k)+\left(\frac{1}{\delta_1}+\frac{1}{\delta_2}+1\right)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\varepsilon^2\\&+\delta_2\tilde{\mathbf{w}}^{\mathrm{T}}(k)\mathbf{s}(\Phi_k)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\mathbf{s}(\Phi_k)^{\mathrm{T}}\tilde{\mathbf{w}}(k)\\=&\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}^{\mathrm{T}}\mathbf{M}_2\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}+\sigma\mathbf{w}^{*\mathrm{T}}\boldsymbol{\Gamma}^{-1}\mathbf{w}^{*}\\&+\frac{\delta_1\delta_2+\delta_1+\delta_2}{\delta_1\delta_2}\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\varepsilon^2,\end{aligned}\tag{18}$$

where

$$\begin{aligned}\mathbf{M}_2=&\begin{bmatrix}\mathbf{M}_{21}&-(\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g+\tilde{\mathbf{C}}_g^{\mathrm{T}})\mathbf{s}^{\mathrm{T}}(\Phi_k)&\tilde{\mathbf{C}}_g^{\mathrm{T}}\sigma\mathbf{s}(\Phi_k)^{\mathrm{T}}\\ *&\mathbf{M}_{22}&0\\ *&*&\mathbf{M}_{23}\end{bmatrix},\\ \mathbf{M}_{21}=&(1+\delta_1)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g-\mathbf{P}+\tilde{\mathbf{C}}_g^{\mathrm{T}}\mathbf{s}^{\mathrm{T}}(\Phi_k)\boldsymbol{\Gamma}\mathbf{s}(\Phi_k)\tilde{\mathbf{C}}_g,\\ \mathbf{M}_{22}=&(1+\delta_2)\mathbf{s}(\Phi_k)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\mathbf{s}(\Phi_k)^{\mathrm{T}}-\sigma\boldsymbol{\Gamma}^{-1},\\ \mathbf{M}_{23}=&(\sigma^2-\sigma)\boldsymbol{\Gamma}^{-1}.\end{aligned}\tag{19}$$

As the next step, a zero element $-\delta_3V(k)+\delta_3\mathbf{x}_g^{\mathrm{T}}(k)\mathbf{P}\mathbf{x}_g(k)+\delta_3\tilde{\mathbf{w}}(k)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{w}}(k)$ is added to the right-hand side of (18) assuming $0<\delta_3<\sigma$. The terms $\delta_3\mathbf{x}_g^{\mathrm{T}}(k)\mathbf{P}\mathbf{x}_g(k)$ and $\delta_3\tilde{\mathbf{w}}(k)^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{w}}(k)$ are included into $\mathbf{M}_{21}$ and $\mathbf{M}_{22}$ of $\mathbf{M}_2$ so that a matrix $\mathbf{M}_3$ is obtained using $\mathbf{s}^{\mathrm{T}}(\Phi)\mathbf{s}(\Phi)\leqslant l$ and $\boldsymbol{\Gamma}\leqslant\lambda_{max}(\boldsymbol{\Gamma})I$:

$$\begin{aligned}\Delta V(k)\leqslant&-\delta_3V(k)+\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}^{\mathrm{T}}\mathbf{M}_3\begin{bmatrix}\mathbf{x}_g(k)\\ \tilde{\mathbf{w}}(k)\\ \hat{\mathbf{w}}(k)\end{bmatrix}\\&+\sigma\mathbf{w}^{*\mathrm{T}}\boldsymbol{\Gamma}^{-1}\mathbf{w}^{*}+\frac{\delta_1\delta_2+\delta_1+\delta_2}{\delta_1\delta_2}\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\varepsilon^2,\end{aligned}\tag{20}$$

where

$$\begin{aligned}\mathbf{M}_3=&\begin{bmatrix}\mathbf{M}_{31}&-(\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g+\tilde{\mathbf{C}}_g^{\mathrm{T}})\mathbf{s}^{\mathrm{T}}(\Phi_k)&\tilde{\mathbf{C}}_g^{\mathrm{T}}\sigma\mathbf{s}(\Phi_k)^{\mathrm{T}}\\ *&\mathbf{M}_{32}&0\\ *&*&\mathbf{M}_{33}\end{bmatrix},\\ \mathbf{M}_{31}=&(1+\delta_1)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g-(1-\delta_3)\mathbf{P}+\lambda_{max}(\boldsymbol{\Gamma})l\tilde{\mathbf{C}}_g^{\mathrm{T}}\tilde{\mathbf{C}}_g,\\ \mathbf{M}_{32}=&(1+\delta_2)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\mathbf{s}(\Phi_k)\mathbf{s}(\Phi_k)^{\mathrm{T}}+(\delta_3-\sigma)\boldsymbol{\Gamma}^{-1},\\ \mathbf{M}_{33}=&(\sigma^2-\sigma)\boldsymbol{\Gamma}^{-1}.\end{aligned}$$

Assuming $\mathbf{M}_3<0$, it follows:

$$\Delta V(k)\leqslant-\delta_3V+\sigma\mathbf{w}^{*\mathrm{T}}\boldsymbol{\Gamma}^{-1}\mathbf{w}^{*}+\frac{\delta_1\delta_2+\delta_1+\delta_2}{\delta_1\delta_2}\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\varepsilon^2\tag{21}$$

which implies ultimate boundedness as the terms $\sigma\mathbf{w}^{*\mathrm{T}}\boldsymbol{\Gamma}^{-1}\mathbf{w}^{*}+[(\delta_1\delta_2+\delta_1+\delta_2)/\delta_1\delta_2]\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g\varepsilon^2$ are bounded.

Step 2: It remains to investigate under what condition $\mathbf{M}_3<0$. From applying Schur's Complement (Boyd, El Ghaoui, Feron, & Balakrishnan, 1994) twice, it follows from $\mathbf{s}(\Phi_k)^{\mathrm{T}}\mathbf{s}(\Phi_k)<l$:

$$\mathbf{s}(\Phi_k)\mathbf{s}(\Phi_k)^{\mathrm{T}}<lI.\tag{22}$$

Thus, there is $\mathbf{M}_4 \geqslant \mathbf{M}_3$, where

$$\mathbf{M}_4 = \begin{bmatrix} \mathbf{M}_{41} & -(\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g + \tilde{\mathbf{C}}_g^{\mathrm{T}})\mathbf{s}^{\mathrm{T}}(\Phi_k) & \tilde{\mathbf{C}}_g^{\mathrm{T}}\sigma\mathbf{s}(\Phi_k)^{\mathrm{T}} \\ * & \mathbf{M}_{42} & 0 \\ * & * & \mathbf{M}_{43} \end{bmatrix},$$

$$\mathbf{M}_{41} = (1+\delta_1)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g - (1-\delta_3)\mathbf{P} + \lambda_{max}(\mathbf{\Gamma})l\tilde{\mathbf{C}}_g^{\mathrm{T}}\tilde{\mathbf{C}}_g,$$

$$\mathbf{M}_{42} = ((1+\delta_2)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g l + (\delta_3 - \sigma)\lambda_{min}(\mathbf{\Gamma}^{-1}))I,$$

$$\mathbf{M}_{43} = (\sigma^2 - \sigma)\mathbf{\Gamma}^{-1}.$$

For $\mathbf{M}_4$ to be negative definite, the principal minors of $\mathbf{M}_4$ have to be negative definite. Thus, from $\mathbf{M}_{41} < 0$, $\mathbf{M}_{42} < 0$ and $\mathbf{M}_{43} < 0$ follows the requirement:

$$(1+\delta_1)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g - (1-\delta_3)\mathbf{P} + \lambda_{max}(\mathbf{\Gamma})l\tilde{\mathbf{C}}_g^{\mathrm{T}}\tilde{\mathbf{C}}_g < 0, \tag{23}$$

$$(1+\delta_2)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g l < (\sigma - \delta_3)\lambda_{min}(\mathbf{\Gamma}^{-1}), \tag{24}$$

$$\sigma < 1. \tag{25}$$

Considering those requirements, Schur's complement implies that $\mathbf{M}_4 < 0$ iff $\mathbf{M}_5 < 0$ for

$$\mathbf{M}_5 = \begin{bmatrix} \mathbf{M}_{51} & \tilde{\mathbf{C}}_g^{\mathrm{T}}\sigma\mathbf{s}(\Phi_k)^{\mathrm{T}} \\ \sigma\mathbf{s}(\Phi_k)\tilde{\mathbf{C}}_g & \mathbf{M}_{52} \end{bmatrix},$$

$$\mathbf{M}_{51} = (1+\delta_1)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g - (1-\delta_3)\mathbf{P} + \lambda_{max}(\mathbf{\Gamma})l\tilde{\mathbf{C}}_g^{\mathrm{T}}\tilde{\mathbf{C}}_g + \frac{(\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g + \tilde{\mathbf{C}}_g^{\mathrm{T}})\mathbf{s}^{\mathrm{T}}(\Phi_k)\mathbf{s}(\Phi_k)(\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g + \tilde{\mathbf{C}}_g)}{\lambda_{min}(\mathbf{\Gamma}^{-1})(\sigma - \delta_3) - (1+\delta_2)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g l},$$

$$\mathbf{M}_{52} = (\sigma^2 - \sigma)\mathbf{\Gamma}^{-1}.$$

It follows $\mathbf{M}_5 \leqslant \mathbf{M}_6$ for

$$\mathbf{M}_6 = \begin{bmatrix} \mathbf{M}_{61} & \tilde{\mathbf{C}}_g^{\mathrm{T}}\sigma\mathbf{s}(\Phi_k)^{\mathrm{T}} \\ \sigma\mathbf{s}(\Phi_k)\tilde{\mathbf{C}}_g & \mathbf{M}_{62} \end{bmatrix},$$

$$\mathbf{M}_{61} = (1+\delta_1)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g - (1-\delta_3)\mathbf{P} + \lambda_{max}(\mathbf{\Gamma})l\tilde{\mathbf{C}}_g^{\mathrm{T}}\tilde{\mathbf{C}}_g + \frac{(\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g + \tilde{\mathbf{C}}_g^{\mathrm{T}})l(\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g + \tilde{\mathbf{C}}_g)}{(\sigma - \delta_3)\lambda_{min}(\mathbf{\Gamma}^{-1}) - (1+\delta_2)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g l},$$

$$\mathbf{M}_{62} = (\sigma^2 - \sigma)\lambda_{min}(\mathbf{\Gamma}^{-1})I.$$

Employing another time Schur's complement, it follows that $\mathbf{M}_3 < 0$ if

$$0 > (1+\delta_1)\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g - (1-\delta_3)\mathbf{P} + \lambda_{max}(\mathbf{\Gamma})l\tilde{\mathbf{C}}_g^{\mathrm{T}}\tilde{\mathbf{C}}_g + \frac{(\mathbf{A}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g + \tilde{\mathbf{C}}_g^{\mathrm{T}})l(\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{A}_g + \tilde{\mathbf{C}}_g)}{(\sigma - \delta_3)\lambda_{min}(\mathbf{\Gamma}^{-1}) - (1+\delta_2)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g l} + \frac{\sigma^2\tilde{\mathbf{C}}_g^{\mathrm{T}}\mathbf{s}(\Phi_k)^{\mathrm{T}}\mathbf{s}(\Phi_k)\tilde{\mathbf{C}}_g}{(\sigma - \sigma^2)\lambda_{min}(\mathbf{\Gamma}^{-1})},$$

$$1 > \sigma,$$

$$0 > (1+\delta_2)\mathbf{b}_g^{\mathrm{T}}\mathbf{P}\mathbf{b}_g l - (\sigma - \delta_3)\lambda_{min}(\mathbf{\Gamma}^{-1}). \tag{26}$$

From $\mathbf{s}(\Phi_k)^{\mathrm{T}}\mathbf{s}(\Phi_k) \leqslant l$, it can be implied that if $\lambda_{max}(\mathbf{\Gamma})$ and σ are small enough and $(\lambda_{min}(\mathbf{\Gamma}^{-1})\sigma)$ large enough then (26) is indeed satisfied and the NN-control system augmented to a linear stabilizing controller is indeed ultimately bounded stable. □

Remark 7. Although it is usual to tune the adaptation algorithm on-line, the stability proof for the combination of a linear plant with an unknown disturbance and an NN-controller creates the possibility for a non-linear optimization of $\mathbf{\Gamma}$ and σ, improving the overall controller dynamics. In particular, the requirement (see explanation before Eq. (20))

$$\begin{bmatrix}\mathbf{x}_g^{\mathrm{T}}(k) & \tilde{\mathbf{w}}^{\mathrm{T}}(k) & \hat{\mathbf{w}}^{\mathrm{T}}(k)\end{bmatrix}\mathbf{M}_2\begin{bmatrix}\mathbf{x}_g^{\mathrm{T}}(k) & \tilde{\mathbf{w}}^{\mathrm{T}}(k) & \hat{\mathbf{w}}^{\mathrm{T}}(k)\end{bmatrix}^{\mathrm{T}} + \delta_3\mathbf{x}_g^{\mathrm{T}}(k)\mathbf{P}\mathbf{x}_g(k) + \delta_3\tilde{\mathbf{w}}(k)^{\mathrm{T}}\mathbf{\Gamma}^{-1}\tilde{\mathbf{w}}(k) < 0 \tag{27}$$

could be expressed together with (22) in a set of non-convex matrix inequality conditions which could be solved for given δ_3 to obtain an improved value for $\mathbf{\Gamma}$ and σ. However, this is an area of future research.

Remark 8. As emphasized in Remark 4, the set of ultimate boundedness is minimized for a small value of $\sigma > 0$ as it allows to decrease the magnitude of the second term on the right-hand side of (21). Furthermore, it is possible to decrease the value of this second term in (21) by choosing the values of $\mathbf{\Gamma}$ large. However, these requirements may violate the stability constraints of (26).

References

Abramovitch, D., & Franklin, G. (2002). A brief history of disk drive control. *IEEE Control Systems Magazine*, *22*(3), 28–42.

Abramovitch, D., Wang, F., & Franklin, G. (1994). Disk drive pivot nonlinearity modeling part i: Frequency domain. In *Proceedings of 1994 ACC* (pp. 2604–2607), Baltimore.

Al Mamun, A., & Ge, S. S. (2005). Precision control of hard disk drives. *IEEE Control Systems Magazine*, *25*(4), 14–19.

Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory*. Philadelphia: SIAM.

Calise, A. J., Yang, B.-J., & Craig, J. I. (2002). An augmenting adaptive approach to control of flexible systems. In *Proceedings of AIAA GNCC*, Monterey.

Campa, G., Sharma, M., Calise, A., & Innocenti, M. (2000). Neural network augmentation of linear controllers with application to underwater vehicles. In *Proceedings of 2000 ACC* (Vol. 26, pp. 75–79), Chicago.

Du, C., Xie, L., Guo, G., Zhang, J., Li, Q., & Hredzak, B. (2006). A generalized KYP lemma based control design and application for 425 kTPI servo track writing. In *Proceedings of 2006 ACC*, Minneapolis.

Dupont, P., Hayward, V., Armstrong, B., & Altpeter, F. (2002). Single state elastoplastic friction models. *IEEE Transactions on Automatic Control*, *47*(5), 787–792.

Ge, S. S., Lee, T. H., & Harris, C. J. (1998). *Adaptive neural network control of robotic manipulators*. Singapore: World Scientific.

Ge, S. S., Lee, T. H., Li, G. Y., & Zhang, J. (2003). Adaptive NN-control for a class of discrete-time nonlinear systems. *International Journal of Control*, *76*(4), 334–354.

Ge, S. S., Lee, T. H., & Ren, S. X. (2001). Adaptive friction compensation of servo mechanisms. *International Journal of Systems Science*, *32*(4), 523–532.

Gong, J. Q., Guo, L., Lee, H. S., & Bin, Y. (2002). Modeling and cancellation of pivot nonlinearity in hard disk drives. *IEEE Transactions on Magnetics*, *38*(5), 3560–3565.

Hensen, R. H. A., van de Molengraft, M. J. G., & Steinbuch, M. (2002). Frequency domain identification of dynamic friction model parameters. *IEEE Transactions on Control Systems Technology*, *10*(2), 191–196.

Herrmann, G., Ge, S. S., & Guo, G. (2005). Practical implementation of a neural network controller in a hard disk drive. *IEEE Transactions on Control Systems Technology*, *13*(1), 146–154.

Herrmann, G., & Guo, G. (2004). HDD dual-stage servo-controller design using a μ-analysis tool. *Control Engineering Practice*, *12*(3), 241–251.

Horowitz, R., Li, Y., Oldham, K., Kon, S., & Huang, X. (2007). Dual-stage servo systems and vibration compensation in computer hard disk drives. *Control Engineering Practice*, *15*(3), 291–305.

Hovakimyan, N., Yang, B.-J., & Calise, A. J. (2002). An adaptive output feedback control methodology for non-minimum phase systems. In *Proceedings of CDC* (pp. 949–954), Las Vegas.

Huang, T., Ding, Y., Weerasooriya, S., & Low, T. S. (1998). Disk drive pivot nonlinearity modeling and compensation through fuzzy logic. *IEEE Transactions on Magnetics*, *34*(1), 30–35.

Jia, Q.-W., Wang, Z.-F., & Wang, F.-C. (2005). Repeatable runout disturbance compensation with a new data collection method for hard disk drive. *IEEE Transactions on Magnetics*, *41*(2), 791–796.

Khalil, H. K. (1992). *Nonlinear systems* (2nd ed.). New York: Macmillan Publishing Company.

Nuij, P., Steinbuch, M., & Bosgra, O. (2008). Measuring the higher order sinusoidal input describing functions of a nonlinear plant operating in feedback. *Control Engineering Practice*, *16*(1), 101–113.

Oboe, R., Beghi, A., Capretta, P., & Soldavini, F. C. (2002). A simulation and control design environment for single-stage and dual-stage hard disk drives. *IEEE-ASME Transactions on Mechatronics*, *43*(3), 161–170.

Peng, K., Chen, B., Cheng, G., & Lee, T. (2005). Modeling and compensation of nonlinearities and friction in a micro hard disk drive servo system with nonlinear feedback control. *IEEE Transactions on Control Systems Technology*, *13*(5), 708–721.

Sri-Jayantha, S. M., Dang, H., Sharma, A., Yoneda, I., Kitazaki, N., & Yamamoto, S. (2001). Truetrack^{TM} servo technology for high TPI disk drives. *IEEE Transactions on Magnetics*, *37*(2), 871–876.

Steck, J. E., Rkhsaz, K., & Shue, S.-P. (1996). Linear and neural network feedback for flight control decoupling. *IEEE Control Systems Magazine*, *16*(4), 22–30.

Wang, J., Ge, S. S., & Lee, T. H. (2001). Adaptive friction compensation for servo mechanisms. In G. Tao, & F. L. Lewis (Eds.), *Adaptive control of nonsmooth dynamic systems* (pp. 211–248). Heidelberg: Springer.

Wang, F., Hurst, T., Abramovitch, D., & Franklin, G. (1994). Disk drive pivot nonlinearity modeling part ii: Time domain. In *Proceedings of 1994 ACC* (pp. 2604–2607), Baltimore.

Yang, B.-J., Calise, A. J., & Craig, J. I. (2003). Adaptive output feedback control with input saturation. In *Proceedings of 2003 ACC* (pp. 1572–1577), Denver.

Yang, H. S., Jeong, J., Park, C. H., & Park, Y.-P. (2001). Identification of contributors to HDD servo errors by measuring PES only. *IEEE Transactions on Magnetics*, *37*(2), 883–887.